

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Bandman, Olga
Lal, Preeti
Corley, Neil C.
Shah, Purvi
- (ii) TITLE OF THE INVENTION: HUMAN S-ADENOSYL-L-METHIONINE
METHYLTRANSFERASE
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: To Be Assigned
 - (B) FILING DATE: Herewith
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Billings, Lucy J.
 - (B) REGISTRATION NUMBER: 36,749
 - (C) REFERENCE/DOCKET NUMBER: PF-0352 US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 415-855-0555
 - (B) TELEFAX: 415-845-4166
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 281 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: THP1PLB01
 - (B) CLONE: 10625

Met 1	Ala	Ser	Arg	Gly 5	Arg	Arg	Pro	Glu	His 10	Gly	Gly	Pro	Pro	Glu 15	Leu
Phe	Tyr	Asp	Glu	Thr	Glu	Ala	Arg	Lys	Tyr	Val	Arg	Asn	Ser	Arg	Met
		20						25				30			
Ile	Asp	Ile	Gln	Thr	Arg	Met	Ala	Gly	Arg	Ala	Leu	Glu	Leu	Leu	Tyr
		35					40					45			
Leu	Pro	Glu	Asn	Lys	Pro	Cys	Tyr	Leu	Leu	Asp	Ile	Gly	Cys	Gly	Thr
	50					55					60				
Gly	Leu	Ser	Gly	Ser	Tyr	Leu	Ser	Asp	Glu	Gly	His	Tyr	Trp	Val	Gly
65					70				75					80	
Leu	Asp	Ile	Ser	Pro	Ala	Met	Leu	Asp	Glu	Ala	Val	Asp	Arg	Glu	Ile
				85				90						95	
Glu	Gly	Asp	Leu	Leu	Leu	Gly	Asp	Met	Gly	Gln	Gly	Ile	Pro	Phe	Lys
			100					105					110		
Pro	Gly	Thr	Phe	Asp	Gly	Cys	Ile	Ser	Ile	Ser	Ala	Val	Gln	Trp	Leu
		115					120					125			
Cys	Asn	Ala	Asn	Lys	Lys	Ser	Glu	Asn	Pro	Ala	Lys	Arg	Leu	Tyr	Cys
	130					135					140				
Phe	Phe	Ala	Ser	Leu	Phe	Ser	Val	Leu	Val	Arg	Gly	Ser	Arg	Ala	Val
145					150					155				160	
Leu	Gln	Leu	Tyr	Pro	Glu	Asn	Ser	Glu	Gln	Leu	Glu	Leu	Ile	Thr	Thr
				165					170					175	
Gln	Ala	Thr	Lys	Ala	Gly	Phe	Ser	Gly	Gly	Met	Val	Val	Asp	Tyr	Pro
			180					185					190		
Asn	Ser	Ala	Lys	Ala	Lys	Lys	Phe	Tyr	Leu	Cys	Leu	Phe	Ser	Gly	Pro
		195					200					205			
Ser	Thr	Phe	Ile	Pro	Glu	Gly	Leu	Ser	Glu	Asn	Gln	Asp	Glu	Val	Glu
	210					215					220				
Pro	Arg	Glu	Ser	Val	Phe	Thr	Asn	Glu	Arg	Phe	Pro	Leu	Arg	Met	Ser
				230						235				240	
Arg	Arg	Gly	Met	Val	Arg	Lys	Ser	Arg	Ala	Trp	Val	Leu	Glu	Lys	Lys
				245					250					255	
Glu	Arg	His	Arg	Arg	Gln	Gly	Arg	Glu	Val	Arg	Pro	Asp	Thr	Gln	Tyr
			260					265					270		
Thr	Gly	Arg	Lys	Arg	Lys	Pro	Arg	Phe							
		275					280								

(2) INFORMATION FOR SEO ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1135 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: THP1PLB01
(B) CLONE: 10625

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGTCGCAGGT	GTGCTGCTGA	GCGGTGAGAA	TGGCGTCCCG	CGGCCGGCGT	CCGGAGCATG	60
GCGGACCCCC	AGAGCTGTTT	TATGACGAGA	CAGAAGCCCG	GAAATACTGT	CGCAACTCAC	120
GGATGATTGA	TATCCAGACC	AGGATGGCTG	GCGGAGCATT	GGAGCTTCTT	TATCTGCCAG	180
AGAATAAGCC	CTGTTACCTG	CTGGATATTG	GCTGTGGCAC	TGGGCTGAGT	GGAAGTTATC	240
TGTCAGATGA	AGGGCACTAT	TGGGTGGGCC	TGGATATCAG	CCCTGCCATG	CTGGATGAGG	300
CTGTGGACCG	AGAGATAGAG	GGAGACCTGC	TGCTGGGGGA	TATGGGCCAG	GGCATCCCAT	360
TCAAGCCAGG	CACATTTGAT	GCTTGCATCA	GCATTTCTGC	TGTGCAGTGG	CTCTGTAATG	420
CTAACAAGAA	GTCTGAAAAC	CGTGCCAAGC	GCCTGTACTG	CTTTTCTTGCT	TCTCTTTTTT	480

CTGTTCTCGT	CCGGGGATCC	CGAGCTGTCC	TGCAGCTGTA	CCCTGAGAAC	TCAGAGCAGT	540
TGGAGCTGAT	CACAACCCAG	GCCACAAAGG	CAGGCTTCTC	CGGTGGCATG	GTGGTAGACT	600
ACCCTAACAG	TGCCAAAGCA	AAGAAATTCT	ACCTCTGCTT	GTTTCTGGG	CCTTCGACCT	660
TTATAACCAG	GGGGCTGAGT	GAAAATCAGG	ATGAAGTTGA	ACCCAGGGAG	TCTGTGTTCA	720
CCAATGAGAG	GTTCCCATTA	AGGATGTGCA	GGCGGGGAAT	GGTGAGGAAG	AGTCGGGCAT	780
GGGTGCTGGA	GAAGAAGGAG	CGGCACAGGC	GCCAGGGCAG	GGAAGTCAGA	CCTGACACCC	840
AGTACACCGG	CCGCAAGGCG	AAGCCCCGCT	TCTAAGTCAC	CACGCGGTTT	TGGAAAGGCA	900
CTTGCCCTCTG	CACTTTTCTA	TATTGTTTCT	CTGACAAAGT	AGTATTTTAG	AAAAGTTCTA	960
AAGTTATAAA	AATGTTTTCT	GCAGTAAAAA	AAAAGTTCTC	TGGGCCGGGC	GTGTTGGCTC	1020
ACACCTGTAA	TCCCAGCACC	TTGGGAGGCT	GAGGTGGGAG	GATCATTTGA	GCCAGGAGT	1080
TTGAGACCTG	CCTGGGCAAC	ATAATGAAAC	TTCCTTTCCA	GGGAGAAAAA	AAAAA	1135

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 283 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1065505

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Ala	Ser	Phe	Lys	Val	Lys	Pro	Glu	His	Thr	Gly	Pro	Pro	Asp	Leu
1				5					10					15	
Tyr	Tyr	Asn	Glu	Thr	Glu	Ala	Ala	Lys	Tyr	Ala	Ser	Asn	Ser	His	Ile
		20						25					30		
Thr	Ala	Ile	Gln	His	Glu	Met	Ala	Glu	Arg	Ala	Leu	Glu	Leu	Leu	Ala
		35					40					45			
Leu	Pro	Glu	Gly	Lys	Ser	Gly	Phe	Leu	Leu	Asp	Ile	Gly	Cys	Gly	Thr
	50					55					60				
Gly	Met	Ser	Ser	Glu	Val	Ile	Leu	Asp	Ala	Gly	His	Met	Phe	Val	Gly
65					70				75					80	
Val	Asp	Val	Ser	Arg	Pro	Met	Leu	Glu	Ile	Ala	Arg	Gln	Asp	Glu	Asp
			85						90					95	
Leu	Glu	Ser	Gly	Asp	Phe	Ile	His	Gln	Asp	Met	Gly	Leu	Gly	Met	Pro
			100					105					110		
Phe	Arg	Pro	Gly	Ser	Phe	Asp	Gly	Ala	Ile	Ser	Ile	Ser	Ala	Ile	Gln
		115					120						125		
Trp	Leu	Cys	His	Ala	Asn	Ala	Ser	Asp	Glu	Asn	Pro	Arg	Lys	Arg	Leu
	130					135					140				
Leu	Phe	Phe	Phe	Gln	Ser	Leu	Tyr	Gly	Cys	Leu	Gly	Arg	Gly	Ser	Arg
145					150				155					160	
Ala	Val	Phe	Gln	Phe	Tyr	Pro	Glu	Asn	Asp	Glu	Gln	Cys	Asp	Leu	Ile
			165						170					175	
Met	Gly	Gln	Ala	His	Lys	Ala	Gly	Phe	Asn	Gly	Gly	Leu	Val	Val	Asp
			180					185					190		
Phe	Pro	Glu	Ala	Ala	Lys	Arg	Lys	Lys	Val	Tyr	Leu	Val	Leu	Met	Thr
		195					200					205			
Gly	Gly	Val	Val	Gln	Leu	Pro	Gln	Ala	Leu	Thr	Glu	Asp	Gly	Glu	Glu
	210					215					220				
Ser	Arg	Thr	Gln	Ile	Asp	Asn	Ala	Gly	Arg	Arg	Phe	Val	Trp	Asn	Ser
225					230				235					240	
Arg	Lys	Asn	Glu	Lys	Val	Ala	Lys	Gly	Ser	Lys	Ala	Trp	Ile	Glu	Ala
			245						250					255	
Lys	Arg	Gln	Arg	Gln	Ile	Lys	Gln	Gly	Arg	Asp	Val	Arg	His	Glu	Ser
		260						265					270		

Lys Tyr Ser Gly Arg Lys Arg Lys Thr Lys Phe
275 280

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 275 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1907189

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Ser	Arg	Pro	Glu	Glu	Leu	Ala	Pro	Pro	Glu	Ile	Phe	Tyr	Asn	Asp	1	5	10	15
Ser	Glu	Ala	His	Lys	Tyr	Thr	Gly	Ser	Thr	Arg	Val	Gln	His	Ile	Gln	20	25	30	
Ala	Lys	Met	Thr	Leu	Arg	Ala	Leu	Glu	Leu	Leu	Asn	Leu	Gln	Pro	Cys	35	40	45	
Ser	Phe	Ile	Leu	Asp	Ile	Gly	Cys	Gly	Ser	Gly	Leu	Ser	Gly	Glu	Ile	50	55	60	
Leu	Thr	Gln	Glu	Gly	Asp	His	Val	Trp	Cys	Gly	Leu	Asp	Ile	Ser	Pro	65	70	75	80
Ser	Met	Leu	Ala	Thr	Gly	Leu	Ser	Arg	Glu	Leu	Glu	Gly	Asp	Leu	Met	85	90	95	
Leu	Gln	Asp	Met	Gly	Thr	Gly	Ile	Pro	Phe	Arg	Ala	Gly	Ser	Phe	Asp	100	105	110	
Ala	Ala	Ile	Ser	Ile	Ser	Ala	Ile	Gln	Trp	Leu	Cys	Asn	Ala	Asp	Thr	115	120	125	
Ser	Tyr	Asn	Asp	Pro	Lys	Gln	Arg	Leu	Met	Arg	Phe	Phe	Asn	Thr	Leu	130	135	140	
Tyr	Ala	Ala	Leu	Lys	Lys	Gly	Gly	Lys	Phe	Val	Ala	Gln	Phe	Tyr	Pro	145	150	155	160
Lys	Asn	Asp	Asp	Gln	Val	Asp	Asp	Ile	Leu	Gln	Ser	Ala	Lys	Val	Ala	165	170	175	
Gly	Phe	Ser	Gly	Gly	Leu	Val	Val	Asp	Asp	Pro	Glu	Ser	Lys	Lys	Asn	180	185	190	
Lys	Lys	Tyr	Tyr	Leu	Val	Leu	Ser	Ser	Gly	Ala	Pro	Pro	Gln	Gly	Glu	195	200	205	
Glu	Gln	Val	Asn	Leu	Asp	Gly	Val	Thr	Met	Asp	Glu	Glu	Asn	Val	Asn	210	215	220	
Leu	Lys	Lys	Gln	Leu	Arg	Gln	Arg	Leu	Lys	Gly	Gly	Lys	Asp	Lys	Glu	225	230	235	240
Ser	Ala	Lys	Ser	Phe	Ile	Leu	Arg	Lys	Lys	Glu	Leu	Met	Lys	Arg	Arg	245	250	255	
Gly	Arg	Lys	Val	Ala	Lys	Asp	Ser	Lys	Phe	Thr	Gly	Arg	Lys	Arg	Arg	260	265	270	
His	Arg	Phe														275			